Reviewer's report

Title: Evaluating cutpoints for the Mental Health Inventory (MHI-5) and the Mental Health Component Summary score (MCS) General Health Questionnaire (GHQ-12): A comparison of five different methods

Version: 1 Date: 17 July 2007

Reviewer: Norbert Schmitz

Reviewer's report:

The authors derived cut points for the Mental Health Inventory (MHI-5) and the Mental Health Component Summary score (MCS) using GHQ-12 as a gold standard.

General strengths
This study included a large sample of subjects in the UK. This paper is well written with good methods and correct statistical analysis.

General weaknesses
Though this study has a number of merits, my main concerns are methodological. I have outlined these points below:

Major Points
a) My major concern is the use of the GHQ-12 as gold standard. The GHQ-12 is a screening instrument itself. Therefore, the authors are comparing two screening instruments rather than screening instrument vs gold standard. In my opinion, it is not appropriate to use the GHQ-12 as a gold standard. This is a serious limitation of the study. I would suggest not to focus on cutoff points but to provide information regarding the association between the instruments.

Minor points
b) Dichotomization of the GHQ-12 results in a lost of information. It might be more appropriate to compare several cutoff points for the GHQ-12 or use to Stratum Specific Likelihood Ratio approach as discussed by Furukawa (Furukawa TA, Goldberg DP, Rabe-Hesketh S, et al. Stratum-specific likelihood ratios of two versions of the General Health Questionnaire PSYCHOLOGICAL MEDICINE 31 (3): 519-529 APR 2001).

c) It would be helpful if the authors could provide some information regarding the association between the instruments using a dimensional approach (e.g., correlation, structural equation model)

d) Did the authors control for age and gender, that means was there a similar association for males and females and younger/older subjects?

e) It would be helpful if the authors could provide the wording of the
questionnaires in order to identify overlap.

f) There are different ways to compute a summary score (e.g. Taft C, Karlsson J, Sullivan M. Do SF-36 summary component scores accurately summarize subscale scores? Qual Life Res 2001; 10:395-404). Did the authors use the approach suggested by Ware et al.?

**What next?**: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

**Level of interest**: An article of limited interest

**Quality of written English**: Acceptable

**Statistical review**: No, the manuscript does not need to be seen by a statistician.